Design of a Scalable Nanophotonic Interconnect for Future Multicores

Avinash Kodi and Randy Morris
School of Electrical Engineering and Computer Science
Ohio University, Athens, OH 45701
E-mail: kodi@ohio.edu, rmorris@prime.cs.ohiou.edu

ABSTRACT
As communication-centric computing paradigm gathers momentum due to increased wire delays and excess power dissipation with technology scaling, researchers have focused their attention on developing alternate technology solutions for Network-on-Chip (NoC) architectures. One potential solution is nanophotonics because of higher bandwidth, reduced power dissipation and increased wiring simplification. In this paper, we propose PROPEL, a balanced power and area-efficient on-chip photonic interconnect for future multicores. PROPEL overcomes two fundamental issues facing NoCs architectures, namely power dissipation and area overhead, by a combination of multiplexing techniques (wavelength and space) and by exploiting the recent advances in optical component design space. We also propose a scalable version of PROPEL, called E-PROPEL which can scale to 256 cores. Our results indicate that PROPEL and E-PROPEL are power, cost and area-effective networks when compared to competing on-chip optical topologies when the number of optical components and overall power loss in the network are considered. Simulation results on synthetic traffic indicate that PROPEL performs better (throughput and power) than electrical and optical topologies.

Categories and Subject Descriptors
B.4.3 [Hardware]: Input/Output and Data Communications Interconnections (Subsystems)

Keywords
Optoelectronic, Network-on-Chip, Low-Power architecture, Interconnects

1. INTRODUCTION
Transistor scaling (doubling the number of transistors) every 18 months has enabled computer architects to improve the performance of computing systems. Current trends have shown that performance improvements can be achieved by having many smaller cores that run at reduced supply voltages and frequencies, thereby limiting the power dissipation. However, as hundreds of cores will fit within the same die, non-scalability of bus-based networks combined with the wire delay problem (where the reachable number of transistors within a clock cycle decreases with technology scaling [1]) will limit the performance of future multi-cores. To overcome the combined problems of scalability and wire delay, on-chip wires have moved to more modular and regular Network-on-Chips (NoCs) [2, 3, 4]. However, recent research has also shown that power dissipation, area overhead and network performance are considered major performance bottlenecks for metallic NoCs [5]. In addition, fundamental electrical signaling limitations, electromagnetic interference (EMI), crosstalk and clock skew will cumulatively limit the performance and scalability of electrical interconnects while consuming low power and minimizing the area overhead [6, 7].

Optical technology can provide several significant advantages over metallic interconnects that include: (1) bit rates independent of distance, (2) higher bandwidth due to multiplexing of wavelengths, (3) larger bandwidth density by multiplexing wavelengths on the same waveguide/fiber, (4) lower power by dissipating power only at the endpoints of the communication channel and many more [6]. Optics is the technology of choice at long distances (LAN, WAN) and short distances (board-to-board) as evidenced by industrial products such as Intel Connects, Active Cables and others. The recent surge of nanophotonics components/devices [8] compatible with complementary-metal-oxide semiconductor (CMOS) technology at small footprints, low power and high bit rates have generated significant interest even at the on-chip level [2, 3, 4, 9].

In this paper, we propose PROPEL - a scalable on-chip nanophotonic interconnect that can meet the power and bandwidth demands of future multicores with acceptable optical hardware complexity. We propose PROPEL for 64 cores and in an extended version, called E-PROPEL, we scale the network to 256 cores. It should be noted that PROPEL has been inspired by RAPID architecture and its variation, nD-RAPID [10]. While nD-RAPID design was more suitable for multiprocessor networks, PROPEL incorporates and extends nD-RAPID routing and wavelength assignment (RWA) features for an on-chip implementation. PROPEL differs from nD-RAPID in terms of architectural design (combining cores/processors) and optical implementation (passive network). PROPEL uses optical interconnects for long distance inter-router communication and electrical switching at the routers. This reduces the power dis-
sipation on long inter-router links while electrical switching provides flow control to prevent buffer overflow. We leverage nanophotonic components/devices and exploit optical properties such as wavelength division multiplexing (WDM), space division multiplexing (SDM) and wavelength reuse to reduce power dissipation, increase the bandwidth density and reduce area requirements in an efficient manner. For E-PROPEL, we implement an $N \times N$ arrayed waveguide grating using ring-resonators and silicon waveguides and interconnect, four, 64-core PROPELs to provide scalable inter-chip bandwidth with reduced power consumption. Moreover, we present a detailed optical implementation that includes power and area estimates and performance modeling using network simulation on synthetic traffic traces. Incidentally, PROPEL bears some resemblance to Flattened Butterfly (FB) topology [11] for 64 cores. However, we differ from FB network as we use optical technology for inter-core communication and design a high-performance, energy and area-efficient network. Our results for 64 cores indicate the following: (1) PROPEL reduces the power consumption by 80% when compared to competing electrical networks, (2) PROPEL is comparable and improves performance by more than 10% when compared to electrical and photonic networks with similar bisection bandwidths, and (3) PROPEL requires the least optical hardware (modulators, photodetectors, waveguides) and has the lowest area overhead among competing photonic networks. Moreover, E-PROPEL improves throughput 2-4 times when compared to mesh and reduces power by 60% for 256 cores when compared to electrical networks.

Although there has been considerable work in off-chip optical interconnects, there have been few significant on-chip photonic designs. Collet et al. [12] have concluded that for technology nodes ranging from 0.7 $\mu$m to 0.05 $\mu$m, on-chip lasers will consume the bulk of the power, hindering the design of on-chip photonic networks. Shacham et al. [3] have proposed circuit-switched photonic interconnects, where electronic set-up, photonic communication and teardown have been proposed. The disadvantage of this approach is the excess latency for path set-up which is performed using electrical interconnects. Kirman et al. [9] have proposed an optical bus for intra-chip processor-L2 cache interconnect for 64 cores grouped into 4 sets. However, this design cannot be scaled and bus contention will increase with more cores unless more bandwidth (wavelengths) can be incorporated. Batten et al. [4] have proposed a DRAM-processor interconnect using an opto-electronic global crossbar with electronic arbitration and photonic switching devices using double ring resonators. More recently, CORONA, a 3D-stacked, 256-core, on-chip fully connected optical crossbar with token based optical arbitration has been proposed [2]. This design scales as $O(N^2)$ which increases the cost and complexity of the network.

The significant contributions of this work are as follows: (1) By combining nanophotonics for communication and electronics for switching and arbitration, our work reduces the optical hardware required for designing future photonic networks. It is well known that optical technology is more expensive ($\$) than electronics, therefore our work attempts at reducing the cost-performance ratio. Moreover, this approach is identical to the design of mainstream electrical routers with four cardinal directions which provides natural fault-tolerance, deadlock freedom and flow control. (2) We take on the challenging task of comparing previously proposed nanophotonic architectures and current electrical architectures based on the optical hardware cost, power consumed and performance for synthetic traffic traces. When comparing across different nanophotonic interconnects, we strive hard to choose parameters that are functionally equivalent. (3) We also propose scalable versions of PROPEL for 256 cores and make quantitative comparisons in terms of cost, connectivity and power consumed.

2. SILICON NANOPHOTONIC DEVICES AND COMPONENTS

In this section, we briefly describe the silicon nanophotonic interconnects and components. Nanophotonic interconnect will require (i) lasers to generate the carrier signal, (ii) modulators and drivers to encode the data, (iii) medium (waveguides, fibers, free space) for signal propagation, (iv) photodetectors to detect light and (v) back-end signal processing (transimpedance amplifiers (TIA), voltage amplifiers, clock and data recovery) to recover the transmitted bit.

<table>
<thead>
<tr>
<th>Transmitters</th>
<th>Power (mW/GBps)</th>
<th>Area (mm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VCSEL [13]</td>
<td>2.5</td>
<td>0.36</td>
</tr>
<tr>
<td>Micro-rings [14]</td>
<td>0.01</td>
<td>0.0001</td>
</tr>
<tr>
<td>MZ modulator [15]</td>
<td>5.1</td>
<td>0.025</td>
</tr>
</tbody>
</table>

Transmitters: Table 1 shows three potential transmitters for on-chip applications: vertical-cavity surface emitting lasers (VCSELs), micro-rings resonators, and Mach-Zehnder (MZ) modulators. VCSELs allow for direct modulation of the carrier signal, but consume more power and area and are harder to integrate with III-V semiconductors [16]. Micro-ring resonators and MZ modulators use indirect modulator (external laser) which has the advantage that the power of the laser is not accounted in the total on-chip power budget. While micro-ring resonators are comparable to MZ modulators, micro-ring resonators are more preferred due to their smaller footprint, lower power dissipation and high-speed modulation (12.5 Gbps) [14].

A micro-ring resonator consists of a ring waveguide that is placed in close proximity to an adjacent waveguide. The incoming light traveling along the adjacent waveguide is modulated by changing the refractive index of the ring waveguide. The change in the refractive index of a micro-ring resonator is accomplished through the free-carrier plasma-dispersion effect [8]. In the free-carrier plasma-dispersion effect, an electro-optical process takes place which allows for free carriers to be injected into the ring waveguide when an electric field is applied to it. Accumulation of free carriers changes the refractive index of the ring waveguide allowing for a shift in the coupling wavelength, which will result in light coupling into the ring waveguide. The coupling wavelength of a ring waveguide must satisfy the equation $\lambda_0 \times m = n_{\text{eff}} \times 2\pi R$, where $\lambda_0$ is the wavelength of the incoming optical beam, $m$ is a whole number starting from 1, $n_{\text{eff}}$ is the effective index of the waveguide and $R$ is the radius of the micro-ring. If the equation above is not satisfied by the incoming signal, then it will travel along the adjacent waveguide without coupling into the ring waveguide, thereby providing the
switching activity. Figure 1 shows the two different operations of a micro-ring resonator. In Figure 1(a), the voltage $V_R$ is switched off resulting in the resonant frequency to be different than the incoming signal allowing the light to pass through. In Figure 1(b), the voltage $V_R$ is switched on resulting in the resonant frequency to be equal in the incoming signal allowing the light to be shifted to the drop port. The pre-driver electrical circuit is a chain of tapered inverters used to drive the modulator’s capacitive load (~60fF) [17].

Waveguides: Table 2 shows two potential waveguides for on-chips applications. From the table, Si waveguides have a smaller pitch allowing for higher waveguide density and due to their ease of integration on-chip, Si waveguides are preferred over polymer waveguides for on-chip applications [9]. Recent research into estimating the number of wavelengths that can be multiplexed onto the same waveguide has shown that with single-ring filters and 2.7 GHz free spectral range (FSR), we may have up to 12 wavelengths, which provides around 200 GHz channel spacing between adjacent wavelengths. With double-rings which improves the filtering, it may be possible to pack 64 wavelengths with tighter 60 GHz spacing [4]. Therefore, in our design we utilize 64 wavelengths and extensively reuse these wavelengths to achieve scalable bandwidth.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Si</th>
<th>Polymer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refractive index (μm)</td>
<td>3.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Pitch (μm)</td>
<td>5.5</td>
<td>20</td>
</tr>
<tr>
<td>Time of flight (ps/μm)</td>
<td>10.45</td>
<td>4.93</td>
</tr>
<tr>
<td>Loss (μm)</td>
<td>1.3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2: Various parameters for silicon and polymer waveguides.

Receivers: The optical receiver is composed of light detection (photodiodes), amplification (TIA, voltage amplifier) and clock and data recovery. With the need to absorb light and convert into electrical pulses, Germanium is being used for two reasons: It has significant photo-absorption between 1.1 μm and 1.5 μm and is already being used in CMOS processes [2]. Table 3 shows three different CMOS compatible receiver designs. In this work, we adopt the Si-CMOS [18] optical receiver design due to its low power consumption, area overhead, and 1.1 V operating voltage, when compared to other designs.

3. PROPEL: ARCHITECTURE AND IMPLEMENTATION

In this section, we describe the proposed architecture, PROPEL, its routing and wavelength assignment (RWA), optical implementation and scalability.

Architecture: We choose 22nm technology node for our work as prior research has shown that optics is advantageous compared to electronics at shorter distances [2, 3, 4, 9]. In PROPEL, we combine optical transceivers and electronic switches to design a power-efficient high performance network as shown in Figure 2(a). The proposed off-chip broadband light source will generate $W_N$ wavelengths, $\Lambda = \lambda_0, \lambda_1, \lambda_2, \lambda_3, \ldots \lambda_{WN-1}$. By transmitting the continuous off-chip carrier signal in both x- and y-directions simultaneously, we modulate the signals at the optical transmitters. Figure 2(a) shows 4 cores and a shared L2 cache combined together to form a tile. This grouping reduces the cost of the interconnect as every core does not require lasers attached and more importantly, facilitates local communication through cheaper electronic switching [20]. There are a total of $M$ tiles in the x-direction and $N$ tiles in the y-direction, for a total of $4 \times M \times N$ cores. Each tile is represented by $T(m, n)$ where $0 \leq m \leq M-1$ and $0 \leq n \leq N-1$. With $M = N = 4$, PROPEL can be designed for 64 cores and with $M = N = 8$, PROPEL can be designed for 256 cores. Each tile consists of dual-set (x and y) photonic transceivers and an electronic switch. Figure 2(b) shows the logical connection among the tiles interconnected in two dimensions, x and y. Optical interconnects are used in two dimensions along the grid similar to an electronic 2D mesh or torus. It takes a maximum of 2 hops to traverse any-to-any tile, one hop in the x-dimension and one hop in the y-dimension.

Intra-Tile Interconnect: Each tile consists of a set of modulators (transmitters) and a set of photodetectors (receivers) for both x and y directions. With a shared-L2 cache for the four cores, we will need a 7 × 7 crossbar switch for 64 cores; three for x-direction, three for y-direction and one for the shared-L2. With a private-L2 cache, the crossbar size increases to 10 × 10 crossbar. Research has shown that high-radix routers could monotonically reduce the overall cost of network (power, area and latency) [21]. In addition, these
crossbar and buffer elements are designed in lower metal layers leading to lower power and area overhead with technology scaling. The packet, consisting of several flits, undergoes the usual router stages of RC (routing computation), VC (virtual channel) allocation, SA (switch allocation) and ST (switch traversal). We allow flit interleaving in the electrical domain (intra-tile) and packet interleaving in the optical domain to reduce the contention and processing overhead at the receiver as the optical link rate may not match the electrical router data rate. Flow control signaling is tied to packet flows and not individual flits. This requires additional buffering at the transmitter and receiver ports to hold entire packets and to overcome round-trip control flow information. We use on/off signaling implemented using electrical interconnects based on receiver buffer thresholds.

**Inter Tile Interconnect:** We adopt dimension-order routing (DOR) for inter tile communication. The traffic first flows in the x-direction to an intermediate tile and then flows in the y-direction to reach the destination. We explain the routing in a single dimension (x) involving four tiles and similar design can be extended to y-dimension. Figure 3 shows tiles 0 to 3 arranged along the x-direction. Every tile modulates the same wavelength into a different waveguide. Each destination tile is associated with a waveguide called as the home channel. For example, tile T(0,0) has four modulators (ring resonators), all of which are resonant with the wavelength $\lambda_0$. Three $\lambda_0$ transmissions from tile T(0,0) are used to communicate with the other three tiles T(1,0), T(2,0) and T(3,0) on their home channel waveguides. The fourth resonant wavelength will be used to communicate with the memory bank. As shown in Figure 3, the home channel for tile T(0,0) consists of four wavelengths, $\Lambda = \lambda_0 + \lambda_1 + \lambda_2 + \lambda_3$ transmitted by tile T(0,0), T(1,0), T(2,0) and T(3,0) respectively. The wavelength selective filters located at tile T(0,0) will demultiplex all the wavelengths, except for $\lambda_0$ which originates from itself and is intended for the memory. Similarly, the wavelengths, $\lambda_0$ from tile T(0,0), $\lambda_1$ from tile T(1,0), $\lambda_2$ from tile T(2,0) and $\lambda_3$ from tile T(3,0) are combined and these are used to access the memory banks. These are also the same wavelength at which the above tiles will receive data from the memory module. Our goal is to provide a scalable bandwidth to the memory similar to inter tile communication. While we propose to access off-chip memory using photonic network, these functionalities have yet to be implemented in our simulations. Similar wavelength assignment is replicated even in the y-direction for inter-tile communication.

The RWA algorithm designed for inter tile communication involves selective merging of same wavelengths from source tiles into separate home channels for destination tiles. This design maximizes the bandwidth via WDM and re-uses the same wavelengths on different waveguides via SDM. The electronic switching performs localized arbitration for the output optical transmitters within each tile. As the wavelength for the destination tile is fixed, there is no more contention once the local electronic switching is completed. The effective bandwidth of the nanophotonic interconnect, $B = W_N \times W_{gN} \times B_R$, where $W_N$ is the number of wavelengths, $W_{gN}$ is the number of waveguides and $B_R$ is the
effective bit rate of the channel. With $W_N = 1$, $W_{GV} = 1$ and $B_R = 10$ Gbps, we obtain 10 Gbps of inter-tile communication. It could be possible to increase the bit rates beyond 10 Gbps as has been reported [22], however we may need additional equalization circuits that could consume substantial area on the chip. Another approach to increase the bandwidth is to increase the wavelengths or the waveguides. Increasing the waveguides increases the area occupied by the channels and the transmitter/receiver circuitry, where as increasing the wavelengths increases only the transmitter and receiver circuitry. As prior work has shown the feasibility of using 64 wavelengths, we assume similar number of wavelengths for our approach [2, 4]. As we have four tiles, we divide 64 wavelengths among 4 tiles to provide 160 Gbps of inter tile communication bandwidth. Figure 4 shows a possible implementation of PROPEL. The off-chip broadband signal is split for x-dimension communication, y-dimension communication and DRAM memory banks.

**Scaling PROPEL:** PROPEL can be scaled to 256 cores with $M = N = 8$. Under the constraint that we may be limited to 64 wavelengths, the channel bandwidth may reduce to 80 Gbps (8 wavelengths @ 10 Gbps) with similar wavelength assignment as before. To retain identical inter-chip bandwidth and connectivity (diameter of two), we need to (1) scale the crossbar to $15 \times 15$; 7 in the x-dimension, 7 in the y-dimension and 1 for the cores/caches and (2) increase the number of waveguides by two to meet the inter-tile bandwidth as before.

In order to reduce the size of the crossbar, we propose an alternate design, called (extended) E-PROPEL, where we can increase the communication bandwidth without significantly increasing the cost of the network. We utilize an arrayed waveguide grating that can provide an $N \times N$ switching functionality. AWGs are passive elements which can switch several wavelengths, but have dimensions in centimeters. Recently, an AWG functionality had been implemented using ring resonators [23]. In this work, we extend the design to simultaneously switch 16 wavelengths for $4 \times 4$ AWG. In what follows, we briefly explain the AWG functionality and our proposed E-PROPEL.

Figure 5 shows the $4 \times 4$ AWG. Figure 5(a) shows the functionality and Figure 5(b) shows the implementation. Here, the wavelengths are indicated as $\lambda_{(a-b)}^c$, where $a - b$ indicate the wavelength range and $c$ indicates the input port. For example, consider input port 0. All input wavelengths are indicated as $\lambda_{(0-15)}^0$, $\lambda_{(16-31)}^0$, $\lambda_{(32-47)}^0$, and $\lambda_{(48-63)}^0$. After traversing the series of ring resonators, the wavelengths $\lambda_{(0-15)}^0$ arrive at output port 0, $\lambda_{(16-31)}^0$ at output port 3, $\lambda_{(32-47)}^0$ at output port 2 and $\lambda_{(48-63)}^0$ at output port 1. This enables a $1 \times N$ switching functionality. Now consider input port 1. All input wavelengths are indicated as $\lambda_{(1-15)}^1$, $\lambda_{(16-31)}^1$, $\lambda_{(32-47)}^1$ and $\lambda_{(48-63)}^1$. After traversing the series of ring resonators, the wavelengths $\lambda_{(0-15)}^1$ arrive at output port 1, $\lambda_{(16-31)}^0$ at output port 0, $\lambda_{(32-47)}^0$ at output port 3 and $\lambda_{(48-63)}^0$ at output port 2. In similar manner, a $N \times N$ switching functionality can be implemented.

Figure 5(b) shows the implementation of the 4-input 64 wavelength switching AWGs using ring resonators. Each AWG consists of 4 waveguides such that each waveguide intersects the other three waveguides. At each of these waveguide crossings a select wavelength or wavelength range is switched from one waveguide to the other. This is accomplished by utilizing the upper half of the ring resonator to switch one wavelength from one waveguide to another, where as the lower half of the ring resonator will be used for reverse switching. In Figure 5(b), $\lambda_{(32-47)}^0$ is switched at the intersection of waveguide 0 and waveguide 1 and also at the intersection of waveguide 2 and waveguide 3, $\lambda_{(0-15)}^0$ is switched at the intersection of waveguide 0 and waveguide 3 and also at the intersection of waveguide 1 and waveguide 2, $\lambda_{(16-31)}^1$ is switched at the intersection of waveguide 1 and waveguide 3 and lastly $\lambda_{(48-63)}^1$ is switched at the intersection of waveguide 0 and waveguide 2. The above mentioned switching of selected wavelengths at unique waveguide intersection results in a $4 \times 4$ AWG.

The proposed E-PROPEL is shown in Figure 6. We combine four 64-core PROPELs (called a cluster) using AWGs to design a 256 core E-PROPEL. E-PROPEL is designed as a fatree with multiple roots to provide scalable inter-cluster bandwidth. Every tile with similar coordinates $T(x,y)$ on different clusters are connected together with the $4 \times 4$ AWG.
previously designed. For example, tiles T(0,0) on clusters 0, 1, 2 and 3 are connected together with the AWG. Similarly, tiles T(0,1) on clusters 0, 1, 2 and 3 are connected together with another AWG and so on. Therefore, we will require 16 AWGs for connecting all the tiles from different clusters. This scaling increases the crossbar size to $10 \times 10$ with three more connections between clusters. This reduces the diameter of the network to 3; one hop across clusters and two hops within the cluster.

The 4-input 64 wavelengths AWG crossbar area is estimated based on the 4-input 4 wavelength AWG crossbar developed by [23]. The 4-input 4 wavelength AWG crossbar contains three and four sets of micro-ring resonators in the vertical and horizontal directions respectively. The estimated area overhead for each set is $60 \mu m \times 90 \mu m$. As the 4-input 64 wavelengths AWG crossbar consists of 16 micro-rings resonators at each set, the horizontal and vertical lengths will increase by $16 \times$. This causes the 4-input 64 wavelength AWG area to be $1440 \mu m \times 960 \mu m$. As AWG design occupies considerable area, we evaluate a reduced version of the proposed E-PROPEL, called ER-PROPEL where we retain the AWGs at the top and bottom of the cluster. This alternate design trades-off performance with area and increases the diameter of the network to 4. We discuss the detail results in the next section. While the proposed E-PROPEL provides scalable bandwidth, there are other technological challenges in implementing the proposed architecture such as routing signals to and from the chip and area overhead of AWGs, which are beyond the scope of the paper.

4. PERFORMANCE EVALUATION

In this section, we compare the area and optical hardware complexity of PROPEL to photonic interconnects such as the Shared-Bus from Cornell [9], Processor-DRAM interconnect from MIT [4], CORONA from HP [2] and the Circuit-Switch interconnect from Columbia[3]. We further model and simulate PROPEL and compare to both electrical networks such as the Mesh, Concentrated Mesh (CMesh) [20] and Flattened-Butterfly (FB) [11] and optical networks ([2], [3]) for synthetic traffic traces. The Processor-DRAM [4] was not chosen for performance comparison as they are designed for core-memory interconnect, where as PROPEL is designed for inter-core communication. In what follows, we briefly provide power and area estimates for NoC link and router. Then we compare 64 and 256 core versions of PROPEL with competing electrical and optical networks based on optical hardware required and provide simulation results.

4.1 Electrical Power and Area Estimates

For electrical interconnects, we consider wires implemented in semi-global metal layers for inter-router links. The wire capacitances, resistances and device parameters were obtained from International Roadmap for Semiconductors, Berkeley Predictive Technology Models and several recent publications. At 90nm technology node, we obtain a link power of 10.27 mW for a 1 GHz clock and a $V_{dd}$ of 1.2 V for a flit width of 128 bits [24] by considering a power-optimal repeater insertion. At 22nm with a flit size of 128 bits, the power dissipation will be 198 mW considering a 9 GHz link. To reduce the power dissipation at future technology nodes, we reduce the network frequency to 2 Ghz and reduce the power consumption to 44 mW, which is comparable to power values from [25]. In [25], they use an energy of 19 pJ/flit/hop or 38mw/flit/hop given a 2 GHz clock. The area consumed by the wires is the product of the wire pitch, the number of wires, and the wires length. In this work, we calculated the wire area to be ($\sim 0.0422 \ mm^2/mm$) [24].

For on-chip (SRAM cell-array) buffers at 90 nm technology, an SRAM cell has an estimated width of 1.16 $\mu m$ and a height of 0.87 $\mu m$ [26], giving an area of 1.0092 $\mu m^2$. At 22 nm, we estimate the buffer power to be 8.06 mW and occupies an area of 185 $\mu m^2$, which is similar to the power value from [27]. In [27], they use an energy of 61.7 pJ for a 567 bit flit at 45 nm technology. If a 128 bit flit is scaled to 22 nm, the buffer power would be 6.84 mW. A 5 $\times$ 5 matrix crossbar with tri-state buffer connectors [28] is considered for the regular NoC design. The area of the crossbar is estimated by the number of input/output signals that it should accommodate. At 22 nm, we estimate the power value for
a 5 × 5 crossbar to be 8.66 mW and this value is similar to [27]. We believe that the values obtained here for electrical components (buffers, links, crossbars)[25, 24, 26, 27] closely matches to other network designs, giving us confidence in our calculations.

4.2 Area and Optical Hardware Complexity Analysis

In this section, we analytically compare the optical hardware complexity in terms of wavelengths, optical components (splitters/couplers, ring resonators), total optical power budget, opto-electronic power dissipation and opto-electronic area requirements. For all networks, we assume an off-chip laser source and the following losses consistent across all networks [9, 4, 2, 3]: a star splitter loss (L_S) of -3 × (log_2 N) where N is the number of times the waveguide is split, a splitter/coupler loss (L_C) of -3 dB (50% loss of signal), off-chip laser-to-fiber coupling loss (L_LF) of -0.5 dB, off-chip to on-chip fiber-to-waveguide coupling loss (L_FW) of -2 dB, waveguide loss (L_W) of -1.3 dB/cm, bending loss (L_B) of -1 dB, a modulator traversal loss (L_M) of -1 dB, a waveguide crossover loss (L_X) of -0.05 dB and a waveguide-to-receiver loss (L_WR) of -0.5 dB. In addition, it should be mentioned that the optical loss is the maximum potential optical loss in the system.

PROPEL uses a total of 3,072 ring resonators (192 per tile, 96 each for x- and y-directions), 32 silicon waveguides (16 each for x and y-directions) and 1,536 photodetectors (96 per tile), which results in PROPEL having a total optical area of 64.6 mm². In addition, PROPEL is comprised of 16 electrical routers and 1,536 optical receivers (96 per tile), resulting in a total electrical area of 50 mm². PROPEL maximum optical power loss is given by L_S + L_LF + L_FW + 2 × L_M + L_WR + 4 × L_B + 32 × L_X + L_W, where L_S will be -15 dB (-3*4+32) and L_W will be -6.5 dB. This makes the total power loss to be -32.1 dB.

We followed similar network analysis for every other network and Table 4 shows various optical components and losses for scaled versions of 64 cores. Shared-bus was originally designed with four wavelengths and increasing the number of wavelengths will change most parameters. As will be explained later, shared-bus architecture is limited by the crossbar throughput and therefore, any increase in the wavelength will not change the performance. As can be seen, PROPEL reduces the optical hardware complexity while requiring the least number of ring resonators and has the lowest optical power loss. Moreover, PROPEL can be designed with minimum optical and electrical area overhead. PROPEL requires 3.8× less optical area than CORONA and 2.7× less optical area than Shared-Bus and is comparable to circuit-switch architecture. PROPEL requires 3.8× lesser electrical area than CORONA with the assumption of the specific electrical receiver circuitry [18] adopted for this design.

Table 5 shows various optical components and losses of various photonic interconnects for 256 cores. We compare the Processor-DRAM architecture [4] as this was designed for 256 cores. It should be noted that we did not consider the electrical area overhead for the mesh within the Processor-DRAM architecture which is used for inter-core communication within a group. PROPEL and E-PROPEL are designed for core-to-core communication and utilize comparable components and devices. CORONA requires almost 2× and 3.5× of optical area when compared to PROPEL and E-PROPEL. PROPEL and E-PROPEL require almost 3× lesser electrical area than CORONA architecture. E-PROPEL requires lesser nanophotonic components and occupies lesser area than PROPEL for 256 cores. The proposed architectures PROPEL and E-PROPEL provide a balanced architecture that reduces the number of nanophotonic and electronic components to design an area-efficient and cost-efficient on-chip architecture.

4.3 Throughput, Latency and Power

In this subsection, we first describe our simulation methodology and then our results on synthetic traffic traces. We tested our proposed architecture on several traces such as: (1) Uniform Random, and (2) Permutation Patterns, such as Bit-Reversal, Butterfly, Matrix Transpose, Complement and Perfect Shuffle. Cycle accurate simulator were used to eval-
uate the performance of PROPEL and the above mentioned networks. We assume a packet size of 4 flits with the flit size of 128 bits. Identical bisectional bandwidth and buffering for each electrical network was considered. For FB, we assume delays of 1, 2 and 3 cycles to communicate over 1, 2 and 3 routers respectively to account for longer links in a single dimension. For CORONA and PROPEL, we simulate L2 caches with a crossbar connecting the cores to the optical transmitters to improve performance. CORONA provides a channel bandwidth of 2.56 Tbps and bisection bandwidth of 40.96 Tbps. PROPEL provides a channel bandwidth of 160 Gbps and a bisection bandwidth of 5.12 Tbps. To maintain similar bandwidths for circuit-switch, we assume 240 Gbps of optical channel rate. In simulating the circuit-switch architecture, we assumed a 3 cycle delay to setup each switch element. This 3 cycle delay amounts to RC, SA and ST stages of a regular router pipeline.

4.3.1 Simulation Results

**Synthetic Traffic**: Figure 7 shows the throughput and average network latency per packet for uniform traffic. From Figure 7(a), we can see that CORONA outperforms all network due to its enormous channel bandwidth of 2.56 Tbps as compared to PROPEL which provides only 0.16 Tbps, a 16X reduction. However, PROPEL offers only 15% lower throughput than CORONA with a significant reduction in optical hardware and network cost. PROPEL outperforms Mesh and FB by 33% and 10% respectively with identical bisectional bandwidths. While PROPEL outperforms electrical network, the real advantage of PROPEL can be seen in terms of power dissipation as shown in the next plot. PROPEL is better than circuit-switch traffic by over 50% for uniform traffic. The two-fold reason is that we are considering short packets and the traffic is random. This creates more contention in circuit-switch network and it will not be able to amortize the cost of setting-up the circuit. Shared-bus network saturates early due to the traffic build-up at the two overlapped switches at the entry and exit points from the optical network. There are four cores connected to the first switch, and these four sets of four cores connected to the second switch before entering the optical network. All 16 cores will contend to enter into the shared bus using the two level switches. The network load is significant to saturate the bus even at very low loads. Figure 7(b) shows the average network latency for 64 cores.

The throughput for all traffic traces for various networks are shown in Figure 8(a). In the figure, the results are normalized relative to the mesh network, showing the increase in throughput of each network relative to the mesh. PROPEL’s performance is comparable and even better than most electrical networks and is slightly lower than CORONA. Circuit-switch performs better for Butterfly and Perfect Shuffle traffic traces as these communication traces involve less contention. Shared-bus also improves performance with synthetic traffic traces as select source cores communicate with select destination cores which reduces the random nature of uniform traffic traces. As PROPEL reduces the cost of the network, it trades-off performance with network cost and area. We are currently running GEMS/GARNET with SIMICS to validate our results on all Splash-2 benchmarks.

Figure 8(b) shows the normalized power dissipation. In the figure, the results are normalized relative to the mesh. As seen, PROPEL reduces the power by 5X when compared to mesh network. In fact, all nanophotonic networks reduce the power dissipated when compared to electrical networks with reduced frequency. Increasing the frequency will increase the power dissipation for electrical networks and opto-electronic networks such as PROPEL and Shared-Bus. While CORONA and circuit-switch have least power consumption, we do not take into account the buffering required at the end-points. As these are fully optical networks, buffers will be required at the end-points for receiving and transmitting the packets. This is accounted in PROPEL as backpressure from the channel allows more packets to be in the network.

**256-core Results**: Figure 9(a) shows the normalized throughput and Figure 9(b) shows the normalized power dissipation for mesh, PROPEL, E-PROPEL and RE-PROPEL. As we can see, PROPEL outperforms all other networks for 256 cores, but consumes more area and increases the complexity of the switch. E-PROPEL reduces the size of the crossbar with minimal loss in performance. RE-PROPEL reduces the number of AWG crossbars from 16 to 8 and this results in slight loss in performance as packets have to go over
one extra hop to reach the edge tiles. From Figure 9(b), we can see that PROPEL, E-PROPEL and RE-PROPEL consume power in increasing order as they encounter more hops. PROPEL topologies dissipate 60-70% less power than mesh topology for 256 cores.

5. CONCLUSION

In this paper, we tackled the problem of scalable optoelectronic on-chip interconnects to solve the bandwidth and power dissipation problems of future technology nodes. We designed PROPEL architecture for 22nm technology node using a two-dimensional connectivity. The optical complexity analysis clearly showed that PROPEL is significantly cost-efficient than previously proposed on-chip photonic interconnects while delivering comparable performance at reduced power dissipation. Moreover, this architecture has the desirable scalable features identical to a mesh architecture which can be scaled in two dimensions, and provides fault-tolerance due to multi-path connectivity. We also analyzed the scalability of PROPEL architecture and developed an extended, E-PROPEL architecture. E-PROPEL reduces the crossbar radix while delivering scalable performance for 256 cores.

Acknowledgements: This research was partially supported by NSF grants CCR-0538945, ECCS-0725765 and CCF-0953398.

6. REFERENCES

Figure 9: Simulation results showing (a) saturation throughput and (b) power dissipation for different synthetic traffic patterns for 256 cores. With $M = 8$ and $N = 8$, PROPEL is designed for 256 cores, E-PROPEL is designed by using four 64-core PROPELs with 16 AWG crossbars and RE-PROPEL is designed with four 64-core PROPELs and 8 AWG crossbars.


